

2002.12.10 現代日本論演習 II (田中重人)

第7回「平均値と標準偏差」

0. 尺度水準について：復習
1. 尺度水準と分析法
2. 代表値と散布度
3. 平均値と標準偏差
4. SPSS のコマンド
5. 平均値を使うときの注意事項

【尺度水準について】

- 名義尺度
- 順序尺度
- 間隔尺度
- 比率尺度

【尺度水準と分析法】

名義 × 名義 → クロス表

名義 × 間隔 → 平均値の比較

【代表値と散布度】

★ 平均値 (mean) — 標準偏差 (SD)
(間隔尺度以上)

★ 中央値 (median) — 四分位偏差 (Q)
(順序尺度以上)

(教科書 p. 42–51)

【平均値】

総和をデータ数で割ったもの

【標準偏差】

平均値からの偏差の 2 乗値の平均が「分散」
分散の平方根が「標準偏差」

★ 平均値と標準偏差はセットで使う

★次のデータの平均と SD は?

$\{0, 1, 4, 5, 7\}$

【SPSS のコマンド】

「記述統計」 → 「記述統計」

→ 変数とオプションを指定

【平均値を使うときの注意事項】

- ★平均値ははずれ値の影響を受けやすい。
あまりにかけはなれたケースがあるときは
 - 上下数%を取りのぞいたデータセットで計算する (調整平均：教科書 p. 46)
 - 順位に変換したり中央値を使って分析

★平均値・標準偏差は**間隔尺度**以上のデータ
に対してしか意味をもたない。

順序尺度の平均値をとっていいのは

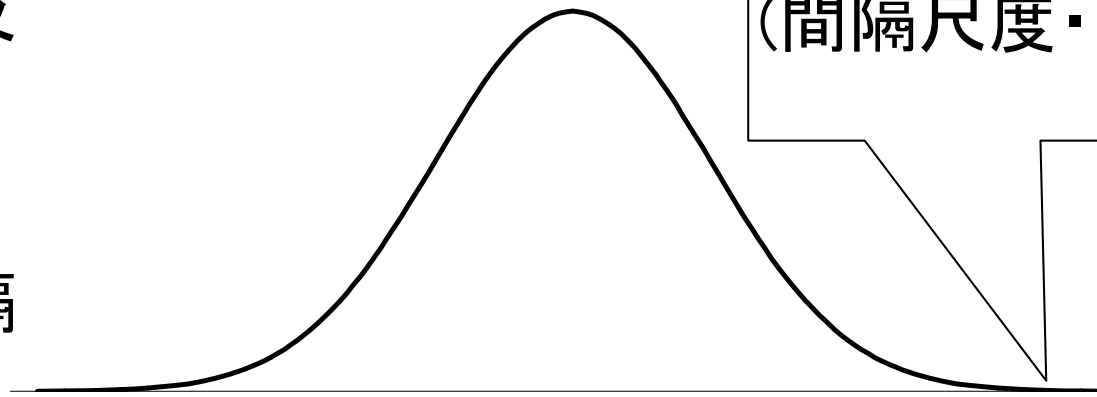
- ・潜在的には間隔尺度のはず
- ・測定のポイントが一定間隔

という2条件をともに満たす場合

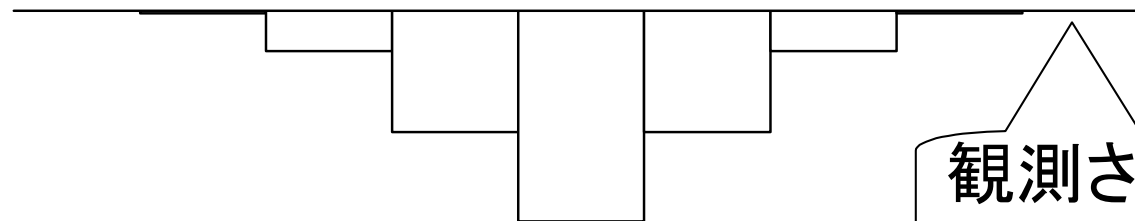
※ 2値の変数は間隔尺度とみなせるが、若干の注意が必要。

観測変数が潜在変数の尺度を反映していると推測できる場合のみ、順序尺度の観測変数を間隔尺度とみなしてよい

潜在変数
(間隔尺度・正規分布)



1 2 3 4 5 6 7 8 9



観測された変数

具体的には

- 4点以上の尺度
- 正規分布に近似 (教科書 p. 53–59) :
 - 単峰性
 - 左右対称性 (歪度)
 - 中央への集中度 (尖度)

ヒストグラムを描いて検討するとよい。

正規分布との乖離度を統計的に検討する手法もある

これらの条件を満たさない場合は

- 非線形変換 (教科書 p.142–144)
- 順位に変換したり中央値を使って分析

※ 間隔尺度のデータでも、
左右対称でないものについては
平均値よりも中央値のほうが
適当であることが多い

典型例：収入・人口など

【課題】

適当な変数について、ヒストグラムの軸上に
平均値と標準偏差を（手書きで）書き入れた
ものを作って提出